The 2002 IAEA intercomparison of software for low-level γ-ray spectrometry*

Dirk Arnold[1], Menno Blaauw[2*], Stjepko Fazinic[3], Vladimir P. Kolotov[4]

[1]Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

[2]Interfaculty Reactor Institute, University of Technology Delft, Mekelweg 15, 2629 JB Delft, The Netherlands

[3]International Atomic Energy Agency, P.O. Box 100, Wagramerstrasse 5, A-1400 Wien, Austria present address Rudjer Boskovic Institute, Bijenicka c. 54, 10000 Zagreb, Croatia

[4]Institute of Geochemistry and Analytical Chemistry of Russian Academy of Sciences, Kosygin str. 19, Moscow B-334, 117975 Russia

The IAEA 2002 set of test spectra for low-level γ-ray spectrometry, reported on in a separate paper, was used in an intercomparison of widely available software packages, i.e. Anges 1.0, GammaVision 5.3, Gamma-W 1.68 for Windows, Ganaas 3.11, Genie2000 2.1, Hyperlab 2002.3.2.18, Interwinner 5.0 and UniSampo 1.97. With each program, efficiency curves were obtained for the two counting geometries (a 500 ml Marinelli beaker on a 33 % relative efficiency HPGe detector, and a 100 ml pillbox on a 96.3 % HPGe detector) and subsequently used to obtain radionuclide activities for the unknown samples. Both the calibration sources and the unknown samples contained radionuclides giving rise to cascade summing effects. Cascade summing correction factors as obtained with some of these programs, as well as with GESPECOR, were compared directly.

After the intercomparison meeting, the activities obtained were compared with the certified activities that had been kept secret until then. In this paper, the results will be presented and suggestions made for further improvement of the software.

1. Introduction

An Advisory Group Meeting (AGM) was organized 5-9 December 1994 in Vienna to discuss the requirements for nuclear analytical software. The meeting mainly addressed questions related to gamma-ray spectrum analysis software. Based on the recommendations from this AGM, an intercomparison of *gamma-ray* spectrum analysis software based on a newly prepared set of test spectra was performed in 1995. The results, including the diskette with the test spectra, were published as an IAEA TECDOC-1011 "Intercomparison of gamma ray analysis software packages". In addition, two articles giving brief explanations of the test spectra, intercomparison method used and the results, were published in this journal [1,2].

The 1995 comparison focused on the ability of the programs to determine peak areas and energies with associated uncertainties. It was concluded that most programs yielded near-optimum results for singlet peaks, that all programs had difficulties with doublets, and that all programs yielded the same peak areas in spite of possible differences in peak shape models used.

However, good peak areas and energies are only the first step in obtaining measurement results in terms of radionuclide activities. When applying relative standardization, the ratio of the peak areas in sample and standard spectrum, applied to the standard's activity, yields the sample's activity, but when one wishes to avoid having to prepare standards for all possible radionuclides, the use of efficiency curves as a function of photon energy and information about gamma-ray yields of the radionuclides is unavoidable. In environmental or low-level γ-ray spectrometry, the determination and application of efficiency curves is hampered by the phenomenon of true coincidence summing, since one tends to bring the sample close to the detector and the probability of the detector registering two photons emitted by the same decaying nucleus as one is no longer negligible [3]. All this becomes even more complex when the sample's specific activity is so low that one needs to measure a larger volume of sample material. Then, the detector efficiency can no longer be considered constant over the source volume and the voluminous true-coincidence summing effect comes into play. Completely neglecting true coincidence summing effects when identifying isotopes may result in negative consequences, such as biased result for radionuclide activities and the associated uncertainties, and missed or falsely identified radionuclides.

* Corresponding author

Solutions to all these problems have been found and presented in the literature in the past, the voluminous effect being the last one to have been resolved [4,5,6,7,8]. Some commercially available programs already have incorporated these recent developments. That is why it was deemed necessary to perform e new intercomparison, which was to resemble the National Physics Laboratory (NPL) intercomparison as reported in August 1997 [9] in many respects.

It was decided at the IAEA that a second set of γ-ray test spectra was required in order to assess the quality of an analysis program for low-level applications, this time consisting of matched sets of calibration and test spectra, with certified activities and detection conditions specified throughout [10]. Also, an intercomparison of currently available software for gamma-ray spectrometry was organized.

In this paper, this intercomparison and its outcome are described.

2. Experimental procedure

2.1. General

Each program was tested during two to three days (including installation of the program) by an expert in the field of gamma-ray spectrometry but without experience with the specific program.

Ganaas failed to deal properly with the calibration spectra, probably due to numbers of counts exceeding the capacity of 16-bit integers, and was excluded from the intercomparison for this reason.

2.2. Library comparison

An efficiency-curve based interpretation of a gamma-ray spectrum hinges on the data in the gamma-ray library. All programs in the intercomparison came with their own libraries. The data in these libraries for selected lines were extracted. Some of these lines are very common, but some are less widely used even though the authors feel they are indispensable. Examples are the ^{109}Cd X-rays, which allow for the efficiency calibration of n-type detectors below 60 keV with a standard calibration source, the 1001 keV line of ^{234m}Pa, which allows for ^{234}Th and ^{238}U determination without extensive sample self-attenuation corrections, and the X-rays of ^{235}U, which are safer to use than the 186 keV, which is interfered by ^{226}Ra.

The data in the program libraries were compared with the “Laboratoire National Henri Becquerel” data from “Nucleide 2000” (version 5 from 2003) as reference data, not because these are considered to be indisputable, but simply to have an independent datasource for comparison purposes. The reference energies were subtracted from the library energies, and the library yields divided by the reference yields. Some programs came with multiple libraries, which were each considered separately.

2.3. Peak area determination

The MIX2NEQ6 spectrum was analyzed with all candidate programs in terms of peak areas and positions, using shape and energy calibrations obtained from the calibration spectra. Peak search thresholds were left at the default value or set to the value advised in the manual of the program. The reported peak areas were compared with the reference peak areas [10] in terms of z-scores and reduced X^2-values, using the same procedures as in the 1995 IAEA intercomparison [2]. Special attention was paid to the small peaks on low background, these being the most important kind in low-level gamma-ray spectrometry.

2.4. Activity determination

2.4.1. Efficiency curve determination

Most of the candidate programs did not claim to have coincidence summing correction abilities. In these cases, the detection efficiency curves were determined specifying only those radionuclides that do not exhibit coincidence summing in the calibration source information, i.e. by omitting ^{134}Cs. The exceptions were GammaVision and Genie 2000, both claiming to allow for coincidence summing effects in the calibration spectrum. As it turned out, Genie 2000 allowed one to tick a box pertaining to this in the calibration step, but this did not affect the outcome in the least. Assuming this to be a simple, soon-to-be-fixed bug, ^{134}Cs was omitted from the calibration source specification for this program as well.

Genie 2000 did require an extra step at this stage: Several spectra of radionuclides emitting only one photon at a time were used to determine the peak-to-total ratio curve as a function of photon energy. (GammaVision extracts this information from the single calibration spectrum. But Genie2000 may use this curve for estimation of true coincidence correction for any measurement geometry, whereas GammaVision requires a special calibration source for each geometry.)

In this way, the SMALL detector was calibrated for all the Marinelli beaker spectra in the set, since all had equal dimensions, density and low-Z matrices. For the pillbox geometry used on the BIG detector, additional corrections had to be applied for the difference in density between the calibration (1.15 g/ml) and the test source (1.6 g/ml). Several programs offered the possibility to somehow do this, the exceptions being Hyperlab and UniSampo. Interwinner merely allowed for the input of a correction curve as a function of gamma-ray energy.

The efficiency curves were considered to be intermediate results in this intercomparison and were not considered to be results of interest.

2.4.2. Coincidence correction factor comparison

Two programs involved in the intercomparison were known to correct for true coincidence summing: GammaVision and Genie 2000. Also, GESPECOR was available to directly calculate coincidence correction factors with. GammaVision normally does not show its correction factors, but the internal engine was available at the intercomparison. Using the efficiency curves as determined with the respective programs themselves for the Marinelli beaker geometry, coincidence correction factors were calculated and tabulated for selected radionuclides. From Genie 2000, such factors could only be extracted from the intermediate results for peaks that were both detected and properly identified.

2.4.3. Radionuclide identification and quantification

All test spectra were analyzed with each program. The activities in the MARITSMALL and PITBIG spectra were converted to the activities at the reference point in time used in the certificates, the activities in the natural radioactivity spectra were reported as observed at the time of measurement.

In many cases, peaks were correctly attributed to the radionuclides and more or less correct activities calculated as intermediate results, but omitted from the final list of radionuclides present in the sample for various reasons. If such an activity could be deduced from the intermediate results written to file by the program, that value was used for comparison with the certified values, but, at the same time, if a radionuclide known to be detectibly present in the source was not reported by a program in the *final* list of results, the result was counted as a miss. Numbers derived directly from such activities are tagged in the tables with asterisks. No attempt was made to express misses in terms of z-scores: The statistics in the spectra were consistently excellent and any miss is a major shortcoming of the program involved, leaving one or more peaks with excellent statistics unexplained.

Similarly, any radionuclide reported but known not to be detectebly present in the sample was counted as a “false hit”. These were expressed in z-scores, since a false hit with a high uncertainty is allowed for in the IAEA intercomparison philosophy. Some effort was made to trace the origin of the false hits with lower uncertainties by studying the intermediate results offered by some of the programs.

Activities reported for radionuclides known to be detectebly present in the sources were counted as “hits”. For each such result, the ratio of reported over certified activity and its uncertainty were calculated, as well as two z-scores: z_{rep} based on the reported and the certified uncertainties and z_{ref} based on certified uncertainties exclusively, as shown in

$$z_{rep} = \frac{A_{rep} - A_{cert}}{\sqrt{s_{rep}^2 + s_{cert}^2}}$$

and

$$z_{ref} = \frac{A_{rep} - A_{cert}}{s_{cert}},$$

where A_{rep} and A_{cert} are the reported and certified activities, and s_{rep} and s_{cert} are the reported and certified 1 standard deviation uncertainties. The first z-score, z_{rep}, could also be calculated for “false hits” by setting A_{cert} and s_{cert} t1o zero. For each spectrum, the z-scores for each category were combined in reduced χ^2 values using

$$\chi_r^2 = \frac{1}{N}\sum_{i=1}^{N} z_i^2 .$$

The χ_r^2 values based on z_{rep} scores should ideally be equal to unity, which would imply that the reported results agree with the certified values to a degree expected from the reported and certified uncertainties. The χ_r^2 values based on z_{ref} scores are expected to be much higher, but reflect the quality of the actual reported activities: The lower, the better therefore. This approach turned out to yield sufficient insight of the performance of the programs for the MARITSMALL and PITBIG spectra. However, the reported activities for the natural radioactivity spectra varied so much in so many respects that a different procedure was applied.

2.5. Activity determination in the natural radioactivity spectra

The various programs yielded output at this stage which was quite different in format and reported entities. Some programs bunched ^{214}Bi and ^{214}Pb and reported them as ^{226}Ra daughters, and likewise for ^{228}Ra. To allow for direct comparison of all programs, the weighted averages of the ^{226}Ra daughters (^{214}Bi and ^{214}Pb) and the ^{228}Ra daughters (^{228}Ac, ^{228}Th, ^{224}Ra, ^{220}Rn, ^{212}Pb, ^{212}Bi and ^{208}Tl) were also computed from the reported nuclide activities for the other programs. The activities were then compared to the certified activities in terms of ratios and z-scores, and χ_r^2 values calculated.

3. Results

The results of the library comparison are shown in Table 1. The authors of Anges reported to have used the same nuclear libraries as used in Gamma-W.

With respect to peak area determination, Figure 1 and Figure 2 show the quality of the peak areas reported, as well as the statistical control thereof, for all hits and for small peaks on low background, respectively. Figure 3 shows how many small peaks were reported correctly, and how many false hits were reported.

The results of the coincidence correction factors comparison are shown in Table 2.

The ratios of reported and certified activities for the MARITSMALL and PITBIG spectra are shown in Table 3, along with their uncertainties. The corresponding χ_r^2 -values are shown in the figures. Figure 4 and Figure 5 show the numbers of reported radionuclides, the misses and false hits, as well as the associated χ_r^2 values for the MARITSMALL spectrum. Figure 6 and Figure 7 show the same for the PITBIG spectrum.

The results obtained for the MIX1EQ and MIX1NEQ spectra are shown in Table 4 and Table 5 for the radionuclides present in the samples. The false hits are shown in Table 6. The results obtained for the MIX2 spectra are qualitatively similar and not shown. The interested reader is referred to the IAEA Technical Document to be published.

4. Discussion

4.1. Library comparison

As Table 1 demonstrates, the values found in the various libraries strongly vary from program to program - the yields sometimes even by a factor of two between the libraries that come with a single program. This is a surprising result implying that the users must ascertain that the libraries supplied with the program in use meet the needs of the application. Especially the useful X-rays of radionuclides of interest are worth checking up on.

4.2. Peak area determination

As can be seen in Figure 1, all programs report reasonably good peak areas, that is, the χ_r^2 computed with the reference uncertainties is low. The best performer in this respect is Anges, the worst GammaVision. Most programs exhibit reasonable statistical control, that is, the reported uncertainties agree well with the actual deviations between the reported peak areas and the reference peak areas, as expressed by low χ_r^2 computed with the reported uncertainties. The best performers in this respect are Anges and Gamma-W; the exception is Hyperlab, which systematically underestimates the uncertainties in peak areas by a factor of 4 or more.

Figure 2 shows that the programs perform roughly the same for small peaks on low background. Hyperlab performs much better for these small peaks than for the larger ones, now yielding the best areas with UniSampo and exhibiting reasonable statistical control.

Figure 3 mostly reflects the threshold setting of the peak search algorithm, and the unavoidable compromise between sensitivity for small peaks and the likelihood of false hits occurring. Genie 2000's default threshold setting apparently is high as compared to the others.

4.3. Coincidence correction factor calculation

Coincidence correction factors can be defined only for those peaks that correspond to photon energies actually emitted by the nucleus. In those cases where all three programs yielded correction factors, the values agree quite well, generally speaking. However, for sum peaks, no such factor can be employed. Genie 2000 fails completely there, since it reports nothing for any pure sum peak.

GESPECOR yields values for some of the sum peaks, but not for all. Oddly, it appears that Genie 2000 performs no corrections for any of the natural radioactivity nuclides, since these are missing from its database of decay schemes. This database cannot be edited or supplemented by the user.

4.4. Nuclide identification and activity determination - MARITSMALL and PITBIG spectra

The MARITSMALL test spectrum is the easiest of all: Five radionuclides in a sample counted under the same circumstances as the calibration source in all respects. Still, the results in Figure 4 surprisingly indicate that not all the candidate programs were able to identify the five radionuclides present in the sample. Some failed to report ^{22}Na, possibly because ^{152}Eu also contributes to the 511 keV which is the most intense (but least characteristic) peak of ^{22}Na. ^{154}Eu was also sometimes reported as the source of the 1274 keV peak of ^{22}Na. GammaVision only reported Ba-133 after the X-rays of Ba-133 had been removed from the library. The resulting misses reflect inadequate radionuclide identification algorithms. Hyperlab performed the worst in this respect.

The false hits also reflect something else: Failure to recognize sum peaks due to coincidence summing as such, and absence of functional logical tests that verify whether other peaks of radionuclides suspected to be present have been found. Anges and Gamma-W apparently have such tests built in. Or, alternatively, their gamma-ray libraries are much less extensive than the libraries of UniSampo, Genie 2000 and Interwinner. This sum peak identification problem is to be expected with a program that does not correct for coincidence summing, but one would have expected Genie 2000 to have recognized them. The only program in the intercomparison that demonstrably recognized sum peaks in the MARITSMALL and PITBIG spectra (because it labels these peaks as such in the intermediate results) is GammaVision.

Figure 5 shows that GammaVision yields the best activities, being in statistical control at the same time (unity “rep hits”χ_r^2), but it failed to report ^{22}Na in the final results. UniSampo clearly performs worst for this spectrum. The results in Table 3 reveal that only Genie 2000 and GammaVision performed coincidence summing corrections: All other programs report results that are too low for the nuclides emitting more than one photon at a time, i.e. all of them except ^{51}Cr. However, Genie 2000 apparently is unaware of ^{22}Na also emitting two positrons coincident with the 1274 keV peak, leading to severe coincidence losses and a 1785 keV sum peak. Unexpectedly, the result for ^{51}Cr seem to suffer from the coincidence summing correction approach, since the programs without summing correction do better for that specific nuclide.

The results in Figure 7 and Figure 6 show fewer identification problems - surprisingly since the same radionuclides are present as in the MARITSMALL spectrum - but far worse results. None of the programs appear to correct for sample self-attenuation adequately. The best statistical control is exhibited by Gamma-W, the best activities are reported by Gamma-W and Genie 2000. GammaVision yields the worst activities but compensates for that somewhat with inflated uncertainties indicating a commendable awareness of their weaknesses of the algorithms.

4.5. Nuclide identification and activity determination - MIX1EQ and MIX1NEQ spectra

These results are the most important ones in the intercomparison, given its scope, and of the two, the MIX1EQ is the most useful because secular equilibrium isexpected in that case. The MIX1NEQ spectrum might exhibit non-equilibrium conditions, depending on the emanation properties of the powdered uranium and thorium ores used.

A first glance at the tables of results reveals that most programs fail to report the presence of most radionuclides present in the sample.

The results show hat none of the programs properly dealt with the 1459.2 - 1460 keV doublet of ^{228}Ac and ^{40}K. In all cases, this peak was attributed to ^{40}K which was hardly present at all in the samples. Hyperlab and GammaVision both failed to correct for ^{228}Ac and ^{40}K in the background, due to a slight mismatch in energy calibration that does not appear to have caused the other programs difficulties.

Genie-2000, GammaVision and UniSampo report a host of radionuclides that are not present in the sample, as shown in Table 6. This is due, in part, to unidentified sum peaks (Genie-200 and UniSampo), and in part to insufficient ability to verify the presence of a radionuclide emitting multiple gamma-rays (GammaVision). The GammaVision and Anges false hits reflect noisy peaks (low z scores), the others erroneous identification of clearly present peaks.

The results for ^{214}Bi reflect the coincidence summing problem: Those programs who do not correct for this are expected to report activities that are too low by a factor of about 0.9. This discrepancy is indeed observed between GammaVision and the others. Anges reports a value close to the GammaVision result for unknown reasons, since Anges does not correct for it.

The results for ^{235}U reflect improper interference correction and/or multiplet deconvolution: Since all programs report reasonably good ^{226}Ra daughter values, they should all be able to subtract the contribution of ^{226}Ra from the 186 keV and come up with a good number for the ^{235}U activity represented by the remainder of the peak. Genie-200 and Gamma-W perform the worst in this respect.

The results for ^{210}Pb indicate major problems in extrapolating the efficiency curve from 59 keV down to 46 keV, as well as sample self-attenuation corrections required because of sample matrix differences between the calibration source and the MIX1 sample.

The ^{226}Ra and ^{228}Ra daughters averages are the only results that somewhat resemble reality across the board, but then these results were in fact not reported by the programs themselves, in most cases, but extracted from the reported activities by the authors.

5. Conclusions

All programs tested in this intercomparison were efficiency-curve based, and all failed in some respect to accomplish their task, as shown in Table 7. Therefore, we can only conclude that matrix- and radionuclide-specific calibration, only requiring good peak area determination, still is the best method for accurate determination of activities in highly efficient counting geometries.

Some of the programs tested did not provide functional sample self-attenuation correction mechanisms. Most did not provide coincidence summing corrections. Programs that do not perform these corrections internally can be assisted by an external program like GESPECOR, but that solution is fully effective only if the gamma-ray library can be modified so that its corrections have their effect at the radionuclide identification stage. And then, of course, the analysis program still needs functional isotope identification and interference correction algorithms to make proper use of the data in the library and the spectrum.

There is hope, however: Each step in the efficiency-curve based approach is performed well by one or more of the participating programs, so potentially, a program that could perform the task well could become available in the near future.

6. References

1. M. Blaauw,V. Osorio Fernandez, W. Westmeier, Nucl.Instr.Meth. A387 (1997) 410
2. M. Blaauw, V. Osorio Fernandez, P. van Espen, G. Bernasconi, R. Capote Noy, H. Manh Dung, N.I. Molla, Nucl.Instr.Meth. A387 (1997) 416
3. K. Debertin, U. Schötzig, Nucl. Instr. Meth. 158 (1979) 471
4. V.P.Kolotov, V.V.Atrashkevich, S.Gelsema. J. Radioanal. Nucl. Chem., 210 (1996) 183
5. V.P. Kolotov, M.J. Koskelo, J. Radioanal. Nucl. Chem. 233 (1998) 95
6. O. Sima and D. Arnold, Appl. Radiat. Isot. 53 (2000) 51
7. D. Arnold, O. Sima, J.Radioanal. Nucl. Chem. 248 (2001) 365
8. M. Blaauw, S.J. Gelsema, Nucl.Instr.Meth. A , 505 (2003) 311-315
9. S.A. Woods, J.D. Hemingway. N.E. Bowles, J.L. Makepeace, “Standard Gamma-Ray Spectra for the Comparison of Spectral Analysis Software”, NPL report CIRM 2, NPL England, 1997.
10. J. M. Los Arcos, M. Blaauw, S. Fazinic, V.P. Kolotov, Nucl.Instr.Meth.A (this issue)

Table 1: Gamma-ray library comparison. The authors of Anges reported to have used the Gamma-W library.

	"Lab. Nat. Henri Becquerel"		Genie 2000		Genie 2000		Gamma-W		Gamma-W		Gamma-W		GammaVision		InterWinner		UniSampo		Hyperlab	
			Standard		Master		Standard		Nuclide		Nature									
Nuclide	E (keV)	Yield (%)	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio	Δ (keV)	ratio
^{60}Co	1173.228	99.85	-0.02	1.00	0.00	1.00	0.01	1.00	-0.03	1.00			0.00	1.00	0.01	1.00	0.01	1.00	0.00	1.00
	1332.492	99.9826	-0.01	1.00	0.01	1.00	0.01	1.00	0.01	1.00			0.02	1.00	0.01	1.00	0.01	1.00	0.00	1.00
	2505.692	2x10-6					0.02	1.00											0.00	0.00
^{109}Cd X	21.991	28.6											0.00	1.03						
	22.163	54											see note							
	25.000	14.7											-0.05	0.97						
	25.484	2.7											0.02	0.89						
^{133}Ba	79.614	2.65	0.01	0.96	0.01	0.99	0.00	0.92	-0.01	0.92							0.00	0.99	0.00	0.99
	80.998	32.9	0.00	1.00	-0.01	1.04	0.00	1.00	0.00	1.00					0.00	1.04	0.00	1.04	0.00	1.04
^{137}Cs	661.657	84.99	-0.11	1.00	0.00	1.00	0.00	1.00	-0.06	1.00	0.00	1.00	-0.04	1.00	0.00	1.00	0.00	1.00	0.00	1.00
^{152}Eu	121.782	28.41	0.00	1.00	0.00	1.00	0.00	1.00	0.02	1.00			0.00	1.03	0.00	1.00	0.00	1.00	0.00	1.01
	964.079	14.49	-0.07	0.99	0.05	0.01	0.02	1.00	0.02	1.01			-0.08	1.01	0.05	0.01	0.02	1.01	0.00	1.01
	1408.013	20.84	-0.06	0.99	1.00	1.00	0.00	1.00	-0.01	1.00			0.07	1.02	0.00	1.00	0.00	1.00	-0.01	1.01
^{210}Pb	46.539	4.25					0.00	1.00	-0.04	0.95			-0.02	0.94	0.00	0.96	0.00	1.00	0.00	1.00
^{228}Ac	209.253	3.89	0.03	1.13	0.00	1.00			0.15	1.16	0.00	1.00	0.15	1.17	0.15	1.12	0.00	1.00	0.00	1.00
	338.320	11.27	0.00	1.01	0.00	1.00			0.08	1.06	0.00	1.00	0.08	1.07	0.08	1.02	0.00	1.00	0.00	1.00
	463.004	4.4	0.00	1.00	0.00	1.01					0.01	1.00	0.00	1.05	0.00	1.00	0.00	1.01	0.00	1.00
	911.204	25.8	0.40	1.07	0.00	1.03			-0.10	1.12	0.00	1.00	-0.13	1.12	-0.13	1.08	0.00	1.03	0.00	1.00
^{234m}Pa	1001.030	0.839	0.00	0.70			0.00	1.00	-0.03	0.97	0.00	1.00							0.00	1.00
^{235}U X	89.954	3.56	0.01	0.42	0.00	0.79							0.01	0.42						
	93.351	5.81	0.00	0.43	0.00	0.77							0.00	0.43					2.74	0.02
	105.550	2.08	-0.55	0.48	-0.55	1.01							-0.55	0.48						
	108.970	0.7	0.17	2.14	0.19	2.14							0.17	2.14					0.19	2.20
^{214}Am	59.541	35.78	0.00	1.02	-0.01	1.01	0.00	1.01	-0.04	1.01			0.00	1.02	0.00	1.01	0.00	1.01	0.00	1.01

Table 2: Comparison of coincidence correction factors. The various symbols mean:
***: sum peak, at apparent yield of 1 %, only listed for Cs-134**
miss: no data in coincidence library
N.D.: Not detected, out of channel range of spectrum or below MD
N.I.: Detected in PICBIG or MIX1NEQLONG spectra but not identified
N.C.: Not calculated by GESPECOR

geometry		E	GESPECOR	Genie 2000	GammaVision
pillbox 1.15 g/ml	Cs-134	232.64	N.C.	N.D.	0.767
		242.72	0.731	N.I.	0.773
		326.59	0.679	N.I.	0.715
		475.34	0.752	0.775	0.786
		563.23	0.736	0.757	0.777
		569.23	0.736	0.758	0.776
		604.69	0.83	0.847	0.865
		795.84	0.831	0.851	0.856
		801.93	0.754	0.778	0.788
		1038.56	0.934	0.954	0.934
	*	1080.07	N.C.	N.I.	0.056
		1167.92	1.136	1.143	1.059
	*	1174.00	N.C.	N.I.	0.558
		1365.16	1.26	0.873	1.122
	*	1400.54	5.25	N.I.	1.861
	*	1406.66	N.C.	N.I.	0.143
	*	1643.30	0.082	N.I.	0.032
	*	1968.89	0.355	N.I.	0.136
	Y-88	898.04	0.886	0.885	0.918
		1836.00	0.872	0.886	0.902
		2734.00	8.394	N.D.	4.379
Marinelli 1.0 g/ml	Ac-228	99.45	0.826	miss	0.649
		209.40	0.94	miss	0.905
		270.30	0.905	miss	0.823
		338.40	0.98	miss	0.966
		794.80	0.896	miss	0.761
		911.07	0.972	miss	0.965
		964.60	0.952	miss	0.908
		968.90	0.973	miss	0.966
		1459.20	0.968	miss	0.951
		1587.90	1.004	miss	1.007
	Bi-212	328.00	0.998	miss	0.998
		727.20	0.972	miss	0.949
		785.40	0.917	miss	0.822
		1620.60	1.008	miss	1.015
	Bi-214	609.32	0.922	miss	0.851
		768.36	0.903	miss	0.811
		934.06	0.909	miss	0.809
		1120.29	0.911	miss	0.813
		1238.11	0.915	miss	0.818
		1377.67	1.028	miss	1.059
		1729.62	1.154	miss	1.298
		1764.52	1.003	miss	1.007
		2204.10	1.002	miss	1.005
	Pb-212	238.63	1	miss	1.000
		300.09	0.901	miss	0.907
	Pb-214	53.23	0.892	miss	0.835
		241.98	0.993	miss	0.996
		295.21	1.001	miss	1.001
		351.92	1	miss	0.999
		785.91	0.994	miss	0.999
	Tl-208	277.24	0.856	N.I.	0.694
		510.72	0.854	N.I.	0.694
		583.13	0.903	N.I.	0.809
		763.30	0.884	N.I.	0.735
		860.10	0.971	N.I.	0.902
		2614.66	0.884	N.D.	0.795

Table 3: Results for the MARITSMALL and PITBIG spectra. Results missing from the final results reported by the programs but retrievable from the intermediate results are tagged with a '*'.

	Gamma-W		Genie 2000		GammaVision		InterWinner		Hyperlab		UniSampo		Anges	
	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err
MARITSMALL														
Ba-133	0.92	0.03	0.97	0.02	1.01	0.04	1.05	0.03	0.93	0.02	0.93	0.05	0.93	0.04
Co-60	0.92	0.04	1.01	0.02	0.99	0.04	0.91	0.02	ND	ND	0.87	0.04	0.91	0.11
Cr-51	1.00	0.04	0.96	0.03	1.06	0.04	0.99	0.03	0.99	0.04	1.00	0.20	0.99	0.04
Eu-152	0.90	0.03	0.97	0.02	1.05	0.04	0.91	0.03	0.89	0.02	0.91	0.06	0.93	0.09
Na-22	0.83	0.04	0.82	0.04	0.99*	0.07	ND	ND	ND	ND	0.77	0.05	0.80	0.12
PITBIG														
Ba-133	0.88	0.04	0.91	0.02	1.47	0.05	0.92	0.03	0.81	0.03	ND	ND	0.83	0.04
Co-60	0.90	0.06	0.91	0.02	0.99	0.04	0.81	0.02	0.79	0.05	0.81	0.03	0.81	0.04
Cr-51	0.92	0.07	0.97	0.02	1.56	0.05	0.75	0.03	0.75	0.05	0.76	0.02	0.74	0.04
Eu-152	0.84	0.04	0.94	0.02	1.17	0.05	0.76	0.03	0.79	0.02	0.82	0.04	0.85	0.04
Na-22	0.78	0.08	0.70	0.02	0.87	0.05	ND	ND	0.69	0.07	0.70	0.04	0.70	0.04

Table 4: Results for the MIX1EQ spectrum

	UniSampo		Inter-Winner		Hyperlab		Genie 2000		Gamma-W		Gamma-Vision		Anges	
	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err
K-40			39.13	11.88	133.31	34.48	41.82	13.82	47.66	16.67	126.64	32.71	31.91	12.22
U-238														
Th-234	0.94	0.02	1.42	0.17									1.13	0.10
Pa-234m	1.73	0.13	14.26	1.01			1.94	0.24	1.03	0.10			1.10	0.06
Ra-226	0.83	0.07							0.28	0.40	1.27	0.05	1.94	0.05
Pb-214	1.12	0.01	1.29	0.01			1.05	0.01					1.05	0.04
Bi-214	0.97	0.01	0.96	0.02			0.96	0.01			1.04	0.04	1.03	0.03
Pb-210	0.52	0.03	0.63	0.06	0.42	0.05			0.68	0.22			2.14	2.76
chisqr	150	1×10^4	100	3×10^5			14	1.4×10^4	0.91	1×10^4	4	1×10^3	50	4×10^4
Ra-226 daughters	1.06	0.00	1.18	0.01			1.01	0.01	1.03	0.03	1.04	0.04	1.04	0.02
$\chi^2_{r\ \mathrm{rep,ref}}$	180	340	230	3×10^3			5.6	21	0.6	70	1.1	140	2.5	150
U-235			1.18	0.13	0.89	0.09	2.68	0.05	2.54	0.10	1.14	0.06	1.18	0.07
Th-231	3.36	1.67	10.48	0.35										
Ac-227									2.01	0.98				
Th-227			0.73	0.09										
Rn-219							0.51	0.22						
pb-211	2.80	0.30												
Th-232														
Ac-228	1.01	0.01	0.98	0.02			0.99	0.01			0.95	0.04		
Th-228	0.98	0.17			1.07	0.15								
Ra-224			1.02	0.04							5.43	0.36		
Rn-220			1.10	0.49										
Pb-212														
Bi-212	46.83	0.97	1.00	0.03			0.62	0.01						
Tl-208	39.80	0.59	0.92	0.02										
$\chi^2_{r\ \mathrm{rep,ref}}$	4×10^3	2×10^7	1.4	12	0.03	4	63	100			20	1×10^4		
Ra-228 daughters	1.02	0.01	0.97	0.02	1.07	0.15	0.91	0.01	0.77	0.04	0.99	0.04	1.05	0.04
$\chi^2_{r\ \mathrm{rep,ref}}$	3	3	3	5	0.2	30	40	45	26	300	0.02	0.2	1.4	12

Table 5: Results for the MIX1NEQ spectrum. Note that due to ^{222}Rn emanation, ^{214}Bi and ^{214}Pb results are expected to be low relative to the certified activities.

	Uni-Sampo		Inter-Winner		Hyper-lab		Genie 2000		Gamma-W		Gamma-Vision		Anges	
	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err	ratio	err
K-40			32.57	8.26	124.95	32.01	33.01	8.68	14.45	4.53	106.37	26.61	31.91	12.22
U-238														
Th-234	1.52	0.03	1.60	0.19									1.13	0.10
Pa-234m	2.81	0.03	6.27	0.18			1.69	0.02	1.19	0.06			1.10	0.06
Ra-226	1.86	0.02			1.14	0.11			1.09	0.12	1.37	0.05	1.94	0.05
Pb-214	1.85	0.01	1.17	0.01			1.07	0.01					1.05	0.04
Bi-214	1.60	0.01	0.93	0.02			0.95	0.01			0.81	0.03	1.03	0.03
Pb-210	0.01	0.00	0.71	0.02	0.39	0.03			0.79	0.25			2.14	2.76
chisqr	$3x10^4$	$1x10^5$	$2x10^2$	$5x10^5$	$7x10^1$	$6x10^3$	$2x10^2$	$8x10^3$	2.0	$1x10^3$	18	$3x10^3$	50	$4x10^3$
Ra-226 daughters	1.74	0.01	1.09	0.01			1.00	0.00	0.99	0.02	0.81	0.03	1.04	0.02
$\chi_r^2{}_{rep, ref}$	$5x10^4$	$5x10^4$	90	800			0.4	0.9	0.2	5	50	$4x10^3$	3	150
U-235	1.78	0.03	1.17	0.03	1.05	0.06	1.40	0.02	1.48	0.45	0.96	0.03	1.18	0.07
Th-231	7.55	0.95	13.38	0.24										
Ac-227									1.04	0.07	25.64	0.91		
Th-227	1.78	0.03	0.89	0.03			1.06	0.02						
rn-219							1.02	0.02						
pb-211	2.66	0.03					2.77	0.04						
Th-232													1.05	0.04
Ac-228	0.58	0.01	0.97	0.02			0.90	0.01						
Th-228	0.62	0.04	0.08	0.02	1.12	0.05								
Ra-224			1.16	0.04							1.10	0.04		
Rn-220			1.31	0.22										
Pb-212														
Bi-212	1.29	0.02	0.94	0.03			0.60	0.01						
Tl-208	1.09	0.02	0.86	0.02										
$\chi_r^2{}_{rep, ref}$	200	300	150	700	0.8	10	100	100			0.8	7	0.2	1.5
Ra-228 daughters	0.63	0.01	0.79	0.01	1.12	0.05	0.84	0.01	0.75	0.04	1.10	0.04		
$\chi_r^2{}_{rep, ref}$	750	750	150	250	6	80	130	130	30	350	6	50	5000	5000

Table 6: Z-scores associated with radionuclides erroneously reported to be present in the MIX1EQ spectrum

	UniSampo	InterWinner	Hyperlab	Genie-2000	Gamma-W	GammaVision	Anges
Na-24	29.8						
Mn-54				24.8			
Fe-59							
Zn-65						1.6	
Ge-73m	6.8						
Kr-97	5.8						
Kr-87	6.4						
Sr-87m	5.5						
Nb-94	9.6					1.4	
Ag-111	6.4						
Cd-109	16.9					7.7	
I-126				4.8			
Te-125m	10.2						
Xe-133				8.3			
Cs-137				137.2			2.8
La-140						3.5	
Eu-154						3.2	
Hf-181						1.6	
Ta-182						4.8	
Pu-239						2.7	

Table 7: Partial or complete failure of the participating programs at various stages of the process.

	Gamma-W	Genie 2000	GammaVision	InterWinner	Hyperlab	UniSampo	Anges
peak area determination							
background correction			X		X		
interference correction	X	X	X	X	X	X	X
coincidence summing effects when calibrating	X	X		X	X	X	X
efficiency correction for sample self-attenuation	X		X	X	X	X	X
coincidence summing effects when analyzing	X	X[1]		X	X	X	X
identification of radionuclides		X[2]	X	X		X[2]	

[1]Partial failure due to absence of natural radionuclides in the databases that come with the program
[2]Failure with respect to false hits due to inability to recognize sum peaks

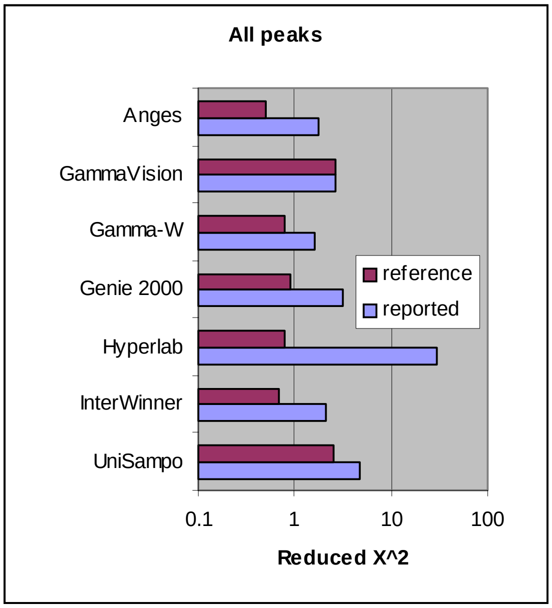

Figure 1: Quality of peak area determination (expressed as the reduced X^2 with reference uncertainties) and statistical control (expressed as the reduced X^2 with reported uncertainties) of the various programs for all peaks in the MIX2NEQ6 spectrum

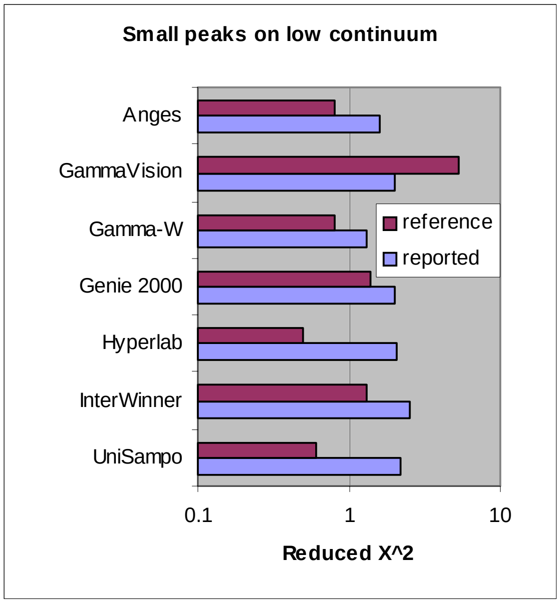

Figure 2: Quality of peak area determination (expressed as the reduced X^2 with reference uncertainties) and statistical control (expressed as the reduced X^2 with reported uncertainties) of the various programs, for small peaks on low background in the MIX2NEQ6 spectrum

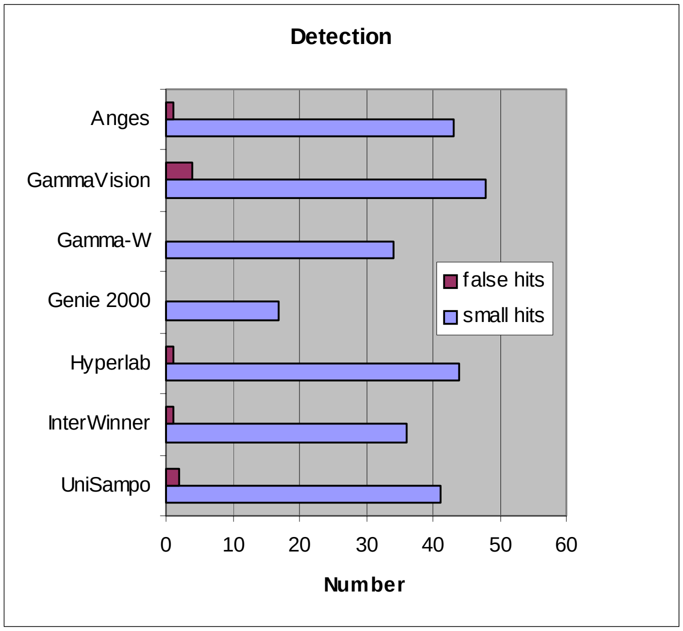

Figure 3: Numbers of small peaks detected and of false hits reported for the MIX2NEQ6 spectrum

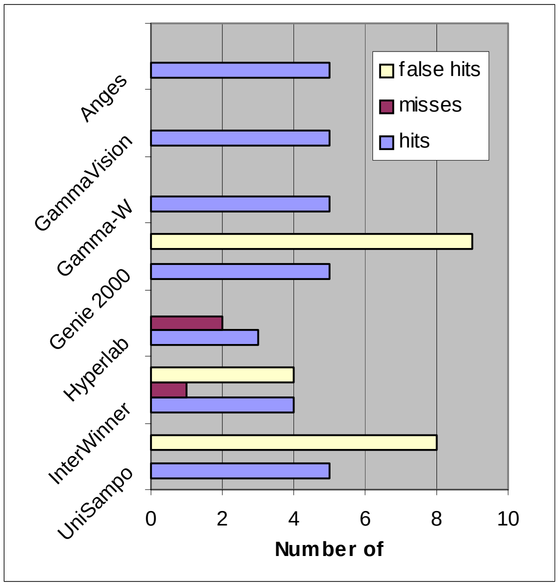

Figure 4: Number of reported radionuclides, misses and false hits in the MARITSMALL spectrum

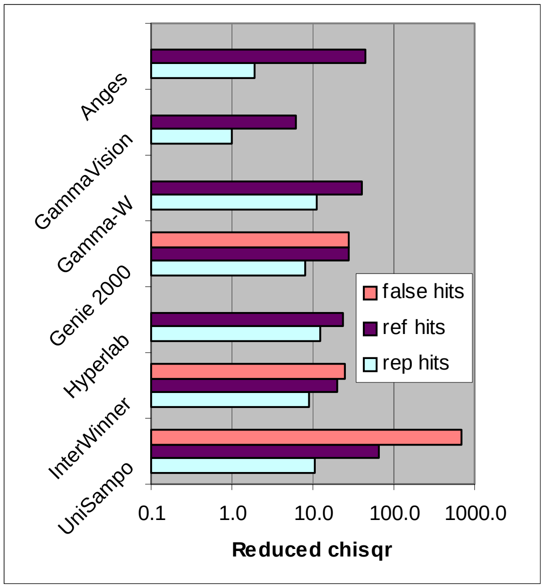

Figure 5: Reduced chisqr-values based on reported uncertainties ("false hits" and "rep hits") and on certified uncertainties ("ref hits") for the MARITSMALL spectrum.

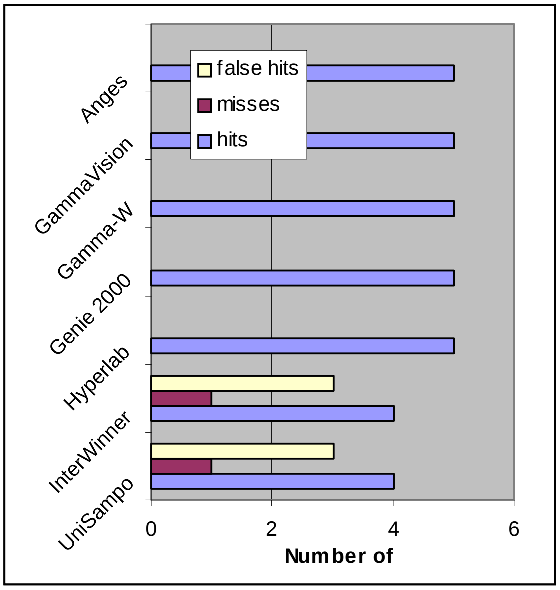

Figure 6: Number of reported radionuclides, misses and false hits in the PITBIG spectrum

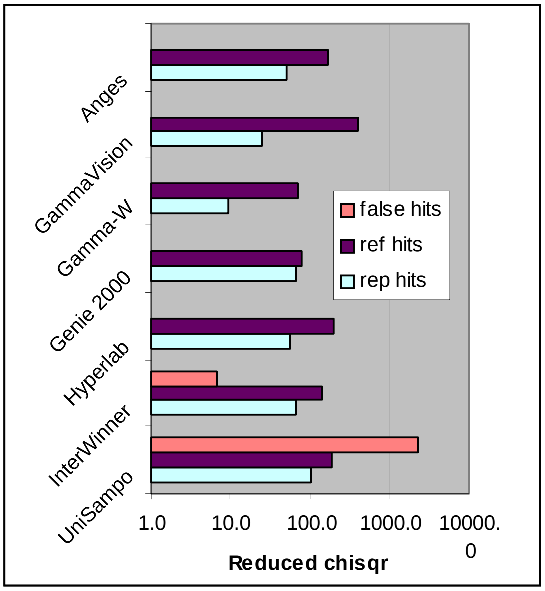

Figure 7: Reduced chisqr-values based on reported uncertainties ("false hits" and "rep hits") and on certified uncertainties ("ref hits") for the PITBIG spectrum.